Using machine learning methods for detecting network anomalies within SNMP-MIB dataset

Ghazi Al-Naymat
Princess Sumaya University for Technology
43 PUBLICATIONS **285** CITATIONS
SEE PROFILE

Mouhammd Alkasassbeh
Princess Sumaya University for Technology
54 PUBLICATIONS **234** CITATIONS
SEE PROFILE

Eshraq Hawari
Mu’tah University
7 PUBLICATIONS **12** CITATIONS
SEE PROFILE

Using machine learning methods for detecting network anomalies within SNMP-MIB dataset

Ghazi Al-Naymat*

Computer Science Department,
Princess Sumaya University for Technology,
Amman, Jordan
Email: g.naymat@psut.edu.jo
*Corresponding author

Mouhammd Al-Kasassbeh
and Eshraq Al-Hawari

Computer Science Department,
Mutah University,
Karak, Jordan
Email: Mouhammd.alkasassbeh@mutah.edu.jo
Email: eshraq.alhawari@gmail.com

Abstract: One of the most prevalent network attacks that threaten networks is Denial of Service (DoS) flooding attacks. Hence, there is a need for effective approaches that can efficiently detect any intrusion in a network. This paper presents an efficient mechanism for network attacks detection within MIB data, which is associated with the protocol (SNMP). This paper investigates the impact of SNMP-MIB data in network anomalies detection. Classification approach is used to build the detection model. This approach presents a comprehensive study on the effectiveness of SNMP-MIB data in detecting different types of attack. The Random Forest classifier achieved the highest accuracy rate with the IP group (100%) and with the Interface group (99.93%). The results show that among five MIB groups the Interface and IP groups are the only groups that are affected the most by all types of attack, while the ICMP, TCP and UDP groups are less affected.

Keywords: anomaly detection; DoS attack; denial of service; SNMP; simple network management protocol; MIB; management information base; machine learning classifier.

Reference to this paper should be made as follows: Al-Naymat, G., Al-Kasassbeh, M. and Al-Hawari, E. (2018) 'Using machine learning methods for detecting network anomalies within SNMP-MIB dataset', *Int. J. Wireless and Mobile Computing*, Vol. 15, No. 1, pp.67–76.

Biographical notes: Ghazi Al-Naymat is an Assistant Professor at Princess Sumaya University for Technology (PSUT), Amman, Jordan. He is currently the Chair of Computer Science Department at King Hussein School of Computing Sciences. He is the founder of the Data Science Master program at PSUT. He received his Master's and PhD degrees from the School of Information Technologies at The University of Sydney, Australia. His research focuses on developing novel data mining techniques for different applications and datasets such as: Graph, Spatial, Spatio-Temporal, and Time Series databases. He is also interested in conducting research in the areas of Big Data. He has published number of papers in excellent international journals and conferences.

Mouhammd Al-Kasassbeh, BSc, MSc, MPhil, PhD, is an Associate Professor at Mutah University specialised in Computer Networks and Security. His research interests include computer networking, sensor network applications, artificial neural network, computer network security, network traffic analysis, network fault detection and abnormality, and time series analysis.

Eshraq Al-Hawari is currently working as a Lecturer at Mutah University, Karak, Jordan. She received her Master's degree in Computer Science from Mutah University, in 2016. Her research interests include wireless sensor networks and network anomalies detection.

This paper is a revised and expanded version of a paper entitled 'Exploiting SNMP-MIB Data to Detect Network Anomalies using Machine Learning Techniques' presented at the 'Intelligent Systems Conference 2018', 6–7 September 2018, London, UK.

1 Introduction

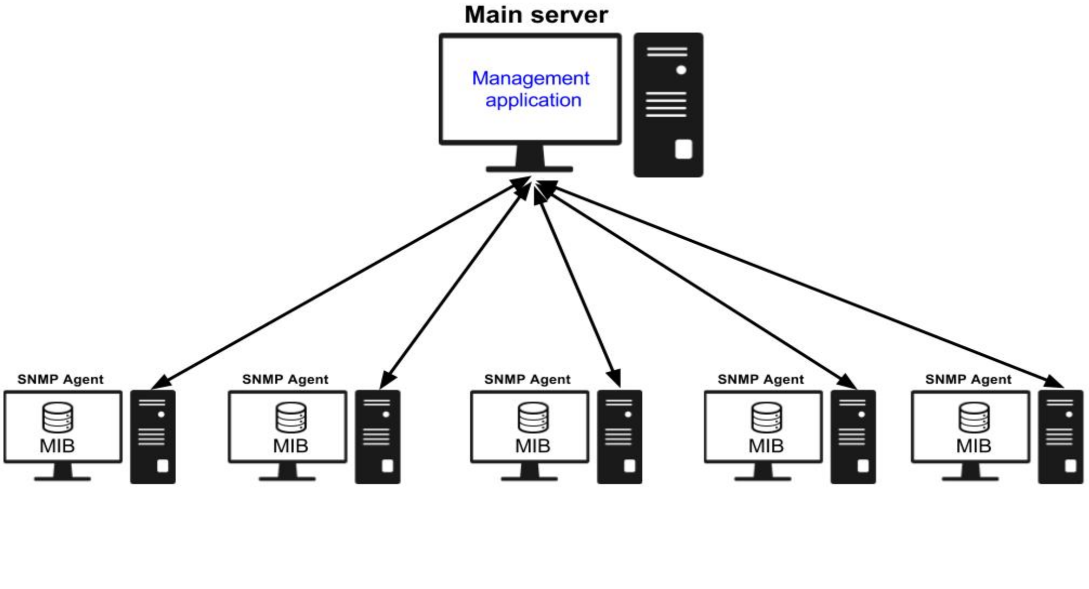
The rapid development of the internet and the growing use of wired and wireless networks have increased the security breaches and malicious attacks. Recently, a variety of network attacks have posed devastating threats to network resources and its security. Some of these attacks are launched to make network services unavailable, such as flooding attacks (e.g., DoS/Distributed Denial of Service (DDoS) and internet worm attacks). Others attacks are used to obtain unauthorised access (e.g., buffer overflow and brute force attacks) (Yu et al., 2008). Among all types of attack, the DoS/DDoS attacks are considered the most significant and dangerous attacks (Alkasassbeh et al., 2016). Moore et al. (2006) reported that these types of attacks are the main threats to the internet, and 90–94% of them are deployed using TCP (Yu et al., 2008). Therefore, it is essential to have rapid detection techniques to detect these attacks with high detection rate for secure and more reliable network services (Kirnapure and Patil, 2017).

Intrusion detection (ID) is a key component of any security mechanism. It aims to protect network security by detecting any abnormal activity occurring in a computer system or network and by comparing it with a normal event to identify signs of intrusion (Nanda and Parikh, 2017). ID is being implemented in two general approaches: Anomaly Intrusion Detection (AID) and Misuse Intrusion Detection (MID) (Gupta and Singh, 2017). AID constructs a normal profile of network behaviour to identify any deviations in normal profile as a possible result of attack activities. AID is based on assumption that the intruder behaviour will be different from the legitimate user behaviour, so it is useful to detect new types of attack. However, MID uses pattern matching to detect intrusions depending on signatures of known attacks to detect it with high accuracy (Bao, 2009).

Intrusion Detection Systems (IDSs) use various data records, collected from target computer or network, to examine them for detecting network attacks (Nanda and Parikh, 2017). A key aspect of any IDS is the type of the used dataset; so obtaining the right type of data about network traffic is essential for accurate and fast intrusion detection (Thottan and Ji, 2003). Most of the current research in the field of network intrusion detection depends on analysing raw packet data to evaluate the security status of computer systems and networks; this will lead to a significant processing burden and late detection time (Yu et al., 2008; Bao 2009). Hence, other data sources about network traffic have been provided by several network management protocols, such as Common Management Information Protocol (CMIP), Remote Network Monitoring (RMON), and Simple Network Management Protocol (SNMP). It should be known that the most deployed and widely used protocol is SNMP Wu and Shao (2005).

SNMP is a popular protocol for network management. It collects information from different types of network devices, such as routers, switches, and hubs on an Internet Protocol (IP) network. SNMP offers powerful statistical information about what is happening in network devices through a database of management variables called Management Information Base (MIB) (Yu et al., 2008; Yu et al., 2013; Wu and Shao, 2005). Figure 1 illustrates that SNMP and MIB data are fundamental elements of the general internet management model. The MIB variables provide numerous traffic informations at different layers and protocols: IP, ICMP, TCP, UDP, etc. The collected information from the network devices can be passively monitored and could be used to characterise network behaviour and, therefore, can be used for network intrusion detection (Gupta and Singh, 2017). By utilising the fine grained data provided from SNMP-MIB in intrusion detection, some of the challenges in network intrusion detection can be avoided, and IDS promises a lower processing overhead analysis with high flexibility of deployment (Yu et al., 2008; Yu et al., 2013; Li and Manikopoulos, 2003). This is what motivates the work presented in this paper, which attempts to present a mechanism for anomaly based intrusion detection using SNMP-MIB dataset.

Figure 1 SNMP-MIB and the general internet management model



Various techniques have been proposed and applied for network anomaly detection using SNMP-MIB data, such as statistical analysis and machine learning methods, which have been extensively used as anomaly detection techniques (Muda et al., 2016). Examples of these approaches include Support Vector Machine (SVM) (Yu et al., 2008; Rahmani et al., 2004), decision trees (Bao, 2009; Namvarasl and Ahmadzadeh, 2014), ANN (Park and Kim, 2008) and *K*-mean clustering (Muda et al., 2016), to mention a few.

The motivation behind this research is set as follows: (1) Due to the importance of IDS in any network environment and the availability of rich SNMP-MIB dataset. (2) Given that some attacks, such as Slowloris and Slowpost, have not been used in previous research. (3) To examine the validity and effectiveness of the use of SNMP-MIB data to classify up to date network attacks.

In this paper, we make the following contributions:

1 We present a MIB based mechanism for network attacks detection and attacks classification using machine learning techniques.

2 We use a recent SNMP-MIB dataset generated by Al-Kasassbeh et al. (2016) to effectively detect DoS and brute force attacks.

3 We use three different classifiers (AdaboostM1, Random Forest and Multilayer Perceptron (MLP) classifiers) for detecting and classifying the DoS attacks (TCP-SYN flooding, UDP flooding, ICMP-ECHO flooding, HTTP flood, Slowloris, Slowpost) and brute force attack.

4 We demonstrate that using SNMP-MIB data with machine learning techniques is a very effective approach for the detection of DoS and brute force attacks.

The rest of the paper is organised as follows: Section 2 includes an overview of related works. Section 3 presents the general architecture of our proposed model for attacks detection and classification based on SNMP-MIB data. Section 4 describes the experiment implementation. In Section 5, we discuss the results. Finally in Section 6, we present our conclusions.

2 Related work

Many researchers have extensively studied anomalies and attacks detection in computer networks during the last decade. Surveys about anomalies detection were presented and many different approaches have been proposed and implemented (Bao, 2009). The vast majority of the solutions presented so far in the scope of network anomalies and attacks detection has been focused on analysing raw traffic features such as number of packets, IP addresses, ports, network flow, etc.). On the other hand, other solutions based on SNMP-MIB data as data source have been proposed for network anomaly detection.

Many studies have exploited SNMP-MIB data in the field of network anomalies detection. Some researchers presented approaches based on statistical analysis of the MIB data, while others have recently utilised machine learning techniques to detect network attacks and other anomalies.

2.1 *Statistical approaches*

The first attempt to exploit SNMP for network security was performed by Cabrera et al. (2002). They proposed a methodology for the early detection of DDoS attacks by applying statistical tests for causality to extract MIB variables that contain precursors to attacks. They used 91 MIB traffic variables from five groups (IP, ICMP, TCP, UDP and SNMP). Three types of DDoS attack (Ping Flood, Targa3 and UDP Flood) were conducted on a research test-bed with controlled loads in traffic. Their work has shown that it is possible to extract a precursor to a DDoS attack using MIB traffic variables and to detect these attacks before the target is shut down with about 1% rate of false alarms.

Al-Kasassbeh and Adda (2009) adopted the distributed model in order to exclude the scalability problems in the network. Their work showed that the statistical methods based on the Wiener filter that upgraded to the mobile agent could be used to detect the abnormality attempts. They took an advantage of the correlation matrix between the input MIB variables and the cross correlation with the desired MIB variables to detect abnormal situations. Al-Kasassbeh used only limited number of MIB variables.

Al-Kasassbeh (2011) applied statistical methods based on the Wiener filter combined with mobile agent technology to detect anomalies in the network traffic by using a set of MIB variables from two MIB groups (Interface and IP). The presented algorithm was tested against four kinds of network attacks (buffer overflow attack, decoy port-scan attack, brute force attack and null session attack) to show the effectiveness of the algorithm in network intrusion detection.

2.2 *Machine learning approaches*

Yu et al. (2008) and Bao (2009) utilised a machine learning approach based on a SVM for network intrusion detection based on SNMP-MIB data. They gathered 13 SNMP-MIB variables corresponding to four MIB groups (IP, ICMP, TCP and UDP). The proposed system was constructed in a hierarchical SVM-based structure for attack traffic detection and classification into different types of attack (TCP-SYN flood, UDP flood and ICMP flood). They concluded that they had achieved fast detection with high detection accuracy 99.27% and with a low rate of false alarms using SVM and the key MIB variables that had been selected from a Correlation Feature Selection (CFS) mechanism. An extended architecture of the system in Yu et al. (2008) and Bao (2009) was proposed by Yu et al. (2013). A system based on C4.5 algorithm with two-level hierarchical structure was performed to detect traffic flooding attacks and classify them into different types of attack (TCP-SYN flood, UDP flood and ICMP flood). Offline, they performed classification and association rule mining facilitated by the C4.5 algorithm, while online they collected SNMP-MIB data and detected DoS/DDoS attacks, subsequently passing the result to the detection module. They reported that attack detection accuracy was about 99.13%.

Hsiao et al. (2009) constructed a detection model based on applying decision tree (C4.5), Naïve Bayesian and SVM data mining techniques that use SNMP-MIB data for Address

Resolution Protocol (ARP) spoofing attack detection. They evaluated the performance of the proposed model using six MIB variables from the IF group, and their results demonstrated that decision tree (C4.5) and SVM produce a better performance by accuracy rate, false alarm rate and missing rate than Naïve Bayesian, while the C4.5 algorithm achieved the highest accuracy rate of about 95.9%.

Namvarasl and Ahmad Zadeh (2014) presented an intrusion detection system based on SNMP-MIB and machine learning. Their system consisted of three modules: the first for selecting key MIB variables from three classification algorithms (C4.5, RIPPER and attribute selection), and the second for generating an intrusion detection model based on the chosen variables and detecting DoS/DDoS attacks in real time in the third module. The dataset used in their system consisted of appropriate MIB variables among 66 variables corresponding to four MIB groups (IP, ICMP, TCP and UDP) and involving a TCP-SYN flood attack, UDP flood attack and ICMP flood attack. Finally, they tested their proposed system and achieved about 99.03% accuracy rate by using a neural network algorithm among three classification algorithms (Neural network, Bayesian network and C4.5).

Cerroni et al. (2013, 2015) presented a decentralised system consisting of a peer-to-peer network of monitoring stations to collect SNMP-MIB statistical data and analyse them using distributed data mining techniques. Cerroni et al. (2013) introduced new supervised distributed classification algorithms (Distributed AdaBoosM1-MultiModel and Distributed AdaBoostM1-SingleModel) for the purpose of network attack detection and classification. The new classification algorithms have been evaluated and tested for decentralised environments using statistical SNMP-MIB data that were gathered locally from each device throughout the experiment. They collected 14 SNMP variables related to IP and TCP protocols involving different types of attack, including DoS, DDoS, TCP port scanning, SSH-DoS and SSH brute force attack. Finally, they concluded that the experimental results of the distributed classification algorithms achieved accuracy in the classification of network attacks.

Cerroni et al. (2015) used unsupervised distributed data clustering techniques based on a *K*-means algorithm provided by the Waikato Environment for Knowledge Analysis (WEKA) tool. They used 14 SNMP variables from the TCP group selected by using the correlation-based feature selection algorithm among SNMP data collected at regular intervals in the experiment. The efficiency of the algorithm was tested for several types of attack, including DoS, DDoS, DoS-SSH and brute force on SSH attack.

Another approach relied on machine learning and data mining techniques for attack detection and classification using SNMP-MIB data was proposed by Priya et al. (2014). They proposed a system named Protocol Independent Detection and Classification (PIDC) to detect and classify Distributed Reflection Denial of Service (DRDoS) attacks, such as a DNS attack and a TCP SYN reflection flooding attack. They captured 13 MIB variables from the TCP and DNS groups, and they used a rank correlation-based detection algorithm to determine the relationship between these variables. The C4.5 classification algorithm was then used with MIB variables to classify the type of attacks that were considered in this research. Their method achieved about 99% true positive rate and 1% false positive rate in the detection of reflected attacks.

Some other studies utilised data obtained from SNMP-MIB for the purpose of intrusion detection in wireless networks (Vyavhare et al., 2012; Puttini et al., 2006). Vyavhare et al. (2012) and Puttini et al. (2006) proposed a multi-agent-based intrusion detection system to detect the intrusion locally in mobile wireless networks using information from SNMP-MIB data. Bayesian classification was used to detect anomalous in network traffic in Mobile Ad Hoc Networks (MANETs) using SNMP-MIB variables.

A key observation drawn from the literature related to SNMP-MIB data-based attack detection is that most of the studies mentioned above are limited to specific types and a limited number of attacks. Also, some of these studies focused on the detection of anomalous traffic as it is distinct from normal without considering the determination of the attack type. So there is still a need to exploit SNMP-MIB as a rich data source for more security in computer networks. The distinction between the previous work and the proposed mechanism is that we intend to offer an effective approach with high detection and accuracy rate compared to others in detecting attacks with SNMP-MIB data. Our work utilises MIB data to classify a greater number of different attacks (TCP-SYN, UDP flood, ICMP-ECHO, HTTP flood, Slowpost, Slowloris and brute force) based on machine learning techniques.

3 Proposed model

In this section, our proposed model is demonstrated and the SNMP-MIB data are described. In addition, a brief description is provided about the used classifiers.

3.1 Design

In this section, we explain the main steps we took in our work for network attack detection and classification using SNMP-MIB data and machine learning techniques. The overall architecture and the main idea of the work model are illustrated in Figure 2. First, a new SNMP-MIB dataset was built from network traffic data collected from a real test-bed network. Second, the collected MIB dataset was used with the feature selection method and the three classifiers, as we stated earlier, are used to build classification models on the MIB data to classify attacks by type.

Towards studying the effect of using MIB data on detecting network attacks, we conducted two sets of experiments: (1) The MIB variables collected were used with an accurate feature selection method to select the most effective variables, and then machine learning classifiers were used to classify the type of attack. (2) The MIB variables collected were categorised into their corresponding groups (Interface, IP, TCP and ICMP), and then the classifiers were applied to each group separately. The remainder of the sections presents our methodology in further detail.

Figure 2 The overall architecture of the detection model

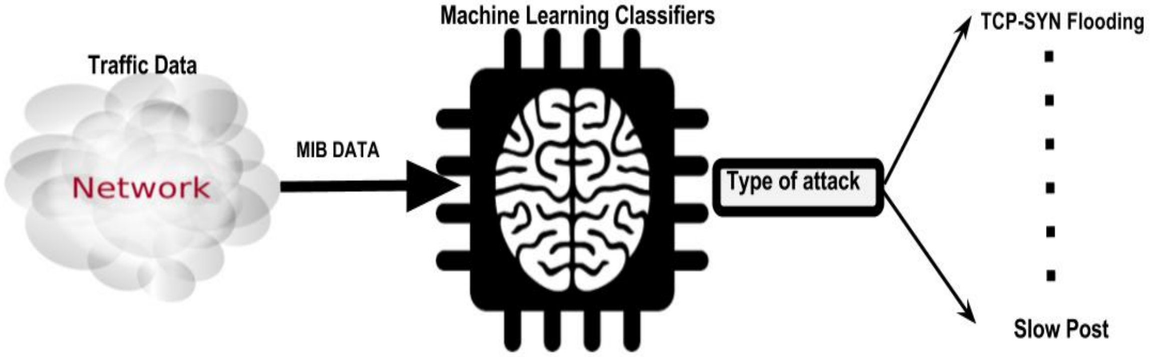


Table 1 34 SNMP-MIB variables

Interface group	*Ip group*	*ICMP group*	*TCP group*	*UDP group*
ifInOctets	ipInReceives	icmpInMsgs	tcpOutRsts	udpInDatagrams
ifOutOctets	ipInDelivers	icmpInDestUnreachs	tcpInSegs	udpOutDatagrams
ifoutDiscards	ipOutRequests	icmpOutMsgs	tcpOutSegs	udpInErrors
ifInUcastPkts	ipOutDiscards	icmpOutDestUnreachs	tcpPassiveOpens	udpNoPorts
ifInNUcastPkts	ipInDiscards	icmpInEchos	tcpRetransSegs	
ifInDiscards	ipForwDatagrams	icmpOutEchoReps	tcpCurrEstab	
ifOutUcastPkts	ipOutNoRoutes		tcpEstabResets	
ifOutNUcastPkts	ipInAddrErrors		tcpActiveOpens	

3.2 SNMP-MIB data

In our experiments, the SNMP-MIB dataset (Al-Kasassbeh et al., 2016) is used for testing our proposed work. The MIB dataset contains approximately 4998 records with 34 MIB variables as shown in Table 1. The data records of attacks fall into six types of DoS attacks (TCP-SYN, UDP flood, ICMP-ECHO, HTTP flood, Slowloris, Slowpost) and brute force attack. Further description for the MIB dataset and attacks instances can be found in Al-Kasassbeh et al. (2016).

3.3 Machine learning classifiers

Classification is one of the most commonly applied supervised machine learning technique. Machine Learning Classifiers have been successfully applied to intrusion detection towards finding various and effective approaches to detect intrusion (Namvarasl and Ahmadzadeh, 2014). Classifiers are used to basically classify the network traffic into normal and abnormal categories. The goal is to build a model from classified objects and use the model to classify new objects as accurately as possible (Yu et al., 2008; Yu et al., 2013). In this work, we have chosen three different classification algorithms: AdaboostM1, Random Forest and MLP. We know from literature that the performance of these classifiers is high in classification problems. These classifiers are applied to our SNMP-MIB dataset in order to classify attack and normal traffic, and then evaluate their accuracy to investigate the ability and effectiveness of SNMP-MIB data in detecting different types of network attacks. To the best of our knowledge, the Random Forest classifier has not been utilised in SNMP-MIB-based attack detection and classification until now, while several classification algorithms, such as BP, C4.5, Bayesian networks and SVM have been used with SNMP-MIB data (Yu et al., 2008; Yu et al., 2013).

4 Experiments

4.1 Implementation

For the purpose of this work, we employed machine learning techniques to evaluate our MIB dataset in attacks detection and classification. For conducting the experiments with classifiers, initially, we randomly divided the MIB dataset into 70% as a training dataset (3498 records) and 30% (1500 records) as a testing dataset. Both datasets contained normal and the other seven attack classes.

In this paper, we applied AdaboostM1, Random Forest and ANN (MLP) classifiers to the MIB datasets. The method categorises the MIB variables in their corresponding groups (Interface, IP, TCP and ICMP groups), and then the classification algorithms are applied to each group separately. Our experiments with machine learning techniques are conducted using an open source data mining toolkit Hall et al. (2009): WEKA 3.7.3. WEKA was originally developed at the University of Waikato, New Zealand. This tool is written in Java language and has a collection of Machine Learning and Data Mining algorithms for data pre-processing, clustering, classification and others (Hall et al., 2009). From WEKA, an ensemble classification method named AdaboostM1 is used with the J48 algorithm as base classifier. Another classifier called Random Forest is also applied with the number of trees equal to 100. The third classifier used is MLP, which is a back propagation learning algorithm, with a learning rate = 0.3, momentum = 0.2 and the number of epochs equal to 500. The transfer function used is Sigmoid function as in equation (1).

$$f(x)=\frac{1}{1+e^{-x}} \tag{1}$$

4.2 Performance evaluation metrics

In our work, we used some important metrics to evaluate how accurately the machine learning classifiers are detecting and classifying the different types of attack based on the MIB dataset. Performance metrics used including Precision, Recall, F-measure and Accuracy as shown in equations (2), (3), (4) and (5). Precision is the ratio of the predicted positive samples that were correctly classified, and Recall is the ratio of positive samples that were classified correctly as positive. It refers to the true positive rate and it is also known as Sensitivity measure. F-Measure is a measure of a classification model's accuracy depending on the precision and the recall metrics. It is considered as a weighted average of both Precision and Recall metrics. Confusion Matrix (Table 2) also is used to evaluate classifier performance. It is a table for visualising the performance of the classification model. It contains information about actual and predicted classifications of the model used.

$$\text{Precision}=\frac{TP}{TP+FP} \tag{2}$$

$$\text{Recall}=\frac{TP}{TP+FN} \tag{3}$$

$$\text{F - Measure}=2\frac{Precision\,Recall}{Precision+Recall} \tag{4}$$

$$\text{Accuracy}=\frac{TP+FN}{TP+FP+TN+FN} \tag{5}$$

Table 2 Confusion matrix for two classes

		Predicted Class	
		Positive	*Negative*
Actual Class	Positive	TP	FP
	Negative	FN	TN

Given that F-Measure is an evaluation measure, which depends on the precision and the recall metrics, it will be considered in all our experimental results as shown in Section 5.

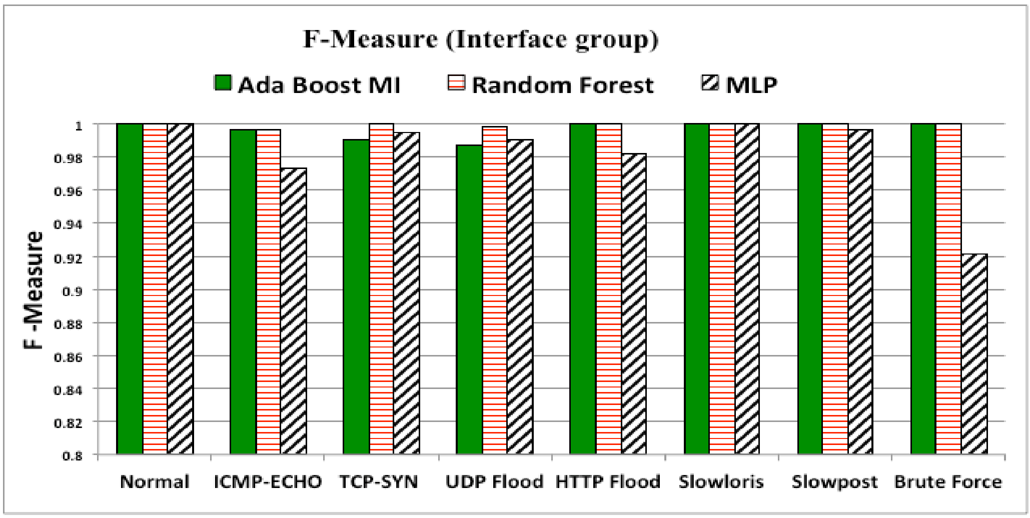
5 Experimental results and discussion

This section describes and explains the experimental results for the proposed method that uses the full MIB dataset (both training and testing datasets). As been mentioned, the MIB data is split into five subsets, where each subset is one MIB group that contains the affiliated MIB variables, and then the classification algorithms (AdaboostM1, Random Forest and MLP classifiers) are used with each group separately. The aim here is to study and show the impact of each group separately in attacks classification by means of the machine learning classifiers.

5.1 Results and discussion of Interface group dataset

In this experiment, the AdaboostM1, Random Forest and MLP classifiers are applied and evaluated on the Interface group dataset. Figure 3 shows the performance of the classifiers in terms F-measure, based on the Interface MIB variables. The results indicate that all classifiers achieved very high results (100%) in identifying the Normal traffic, the Slowloris and Slowpost attack records in the testing set. From the results, AdaboostM1 and Random Forest classifiers succeeded in identifying most types of attack, followed by MLP classifier, which also achieved high performance in attack detection with the interface group.

Figure 3 F-Measure results with interface group



From Figure 3 we can observe that F-Measure results are very high for all classifiers in all types of attack except for the AdaboostM1 classifier in the TCP-SYN and UDP flood attacks where the performance is less effective compared to other attacks. Also, the MLP classifier has the minimum result for the brute force attack, which means that MLP cannot identify brute force attack records among all the other records.

5.2 Results and discussion of IP group dataset

In this experiment, the classifiers are applied to the IP group dataset. Figure 4 shows the performance of the chosen classifiers in terms F-measure, based on IP-MIB variables.

The results presented in Figure 4 demonstrate that all classifiers achieved very high F-measure (100%), except for the MLP classifier showing lower result in the ICMP-ECHO and the brute force attacks. This confirms that most of the attack types and the normal traffic are identified precisely and correctly by all classifiers using the IP group, which means that IP MIB variables are affected by all types of attack.

5.3 Results and discussion of ICMP group dataset

In this experiment, the classifiers are applied to the ICMP group dataset, involving six MIB variables. Figure 5 shows that the F-measure results are different for all classifiers. These results indicate that all classifiers with the ICMP dataset are more effective in detecting HTTP flood and Slowloris attacks.

5.4 Results and discussion of TCP group dataset

In this experiment, the TCP group dataset is used involving six MIB variables. Figure 6 shows that the maximum F-measure results are for all classifiers when identifying Slowloris attack. On the other hand, none of the classifiers managed to detect the brute force attack. This is possibly due to the brute force attack occurring at the network layer and the TCP MIB variables reflect the network traffic statistics at the transport layer, which lies over the network layer, so that TCP MIB variables are not affected by this attack.

Figure 4 F-measure results with IP group

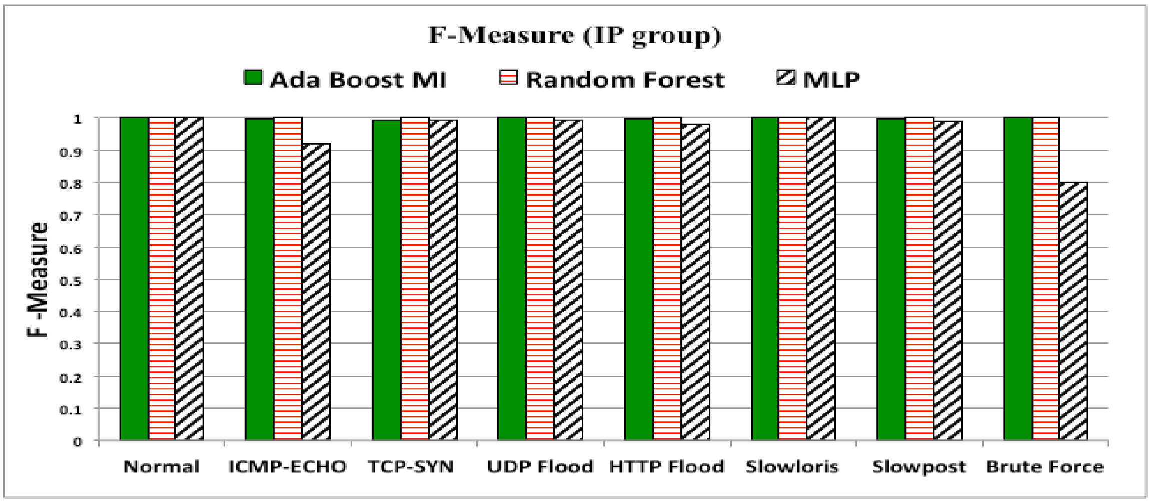


Figure 5 F-measure results with ICMP group

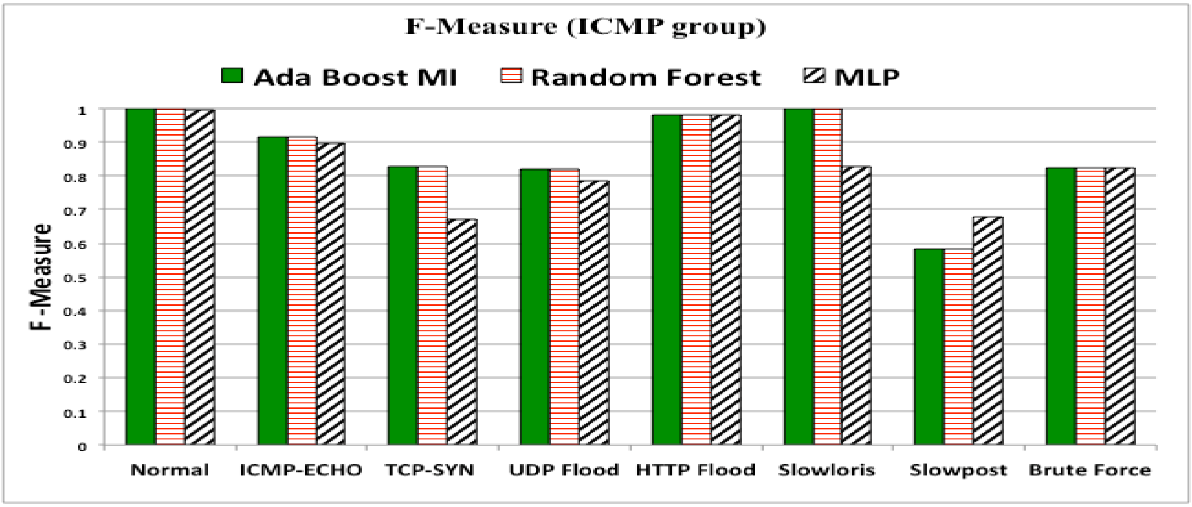
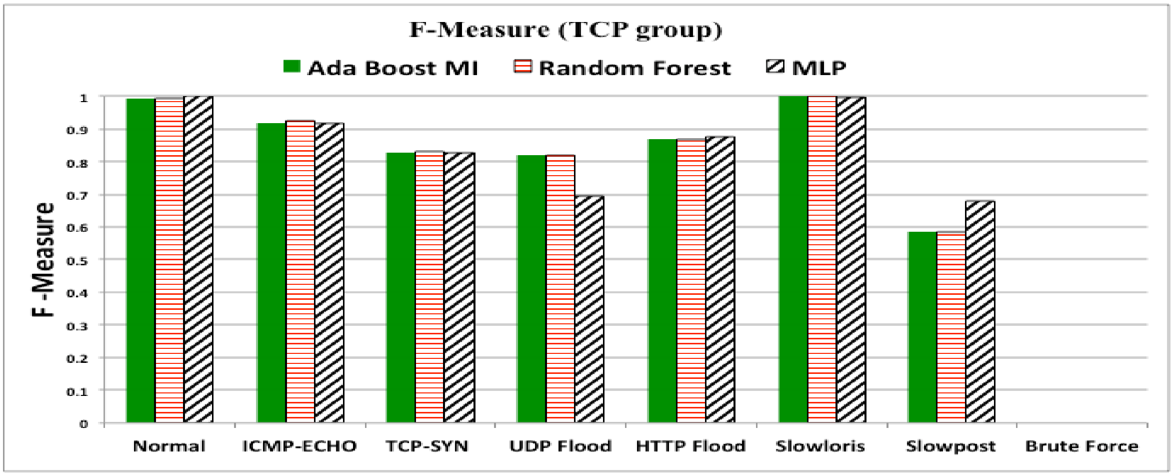


Figure 6 F-measure results with TCP group



5.5 *Results and discussion of UDP group dataset*

In this experiment, the classifiers are applied to the UDP group dataset, involving four MIB variables. The results presented in the Figure 7 demonstrate that all classifiers achieved very high F-measure (100%), except for the MLP classifier, which scores lower results with value of 60% for Slowpost and 83% for the brute force attack. This confirms that the UDP MIB variables have a significant impact on the detection of most types of attack by the three classifiers.

Experimental results in terms of accuracy rate for all classifiers with each MIB group dataset are shown in Figure 8. As depicted, the accuracy results indicate that the Random Forest classifier outperformed in classifying the attacks with all MIB group datasets, followed by the AdaboostM1 classifier. However, the MLP classifier showed the lowest accuracy rates.

Figure 7 F-measure results with UDP group

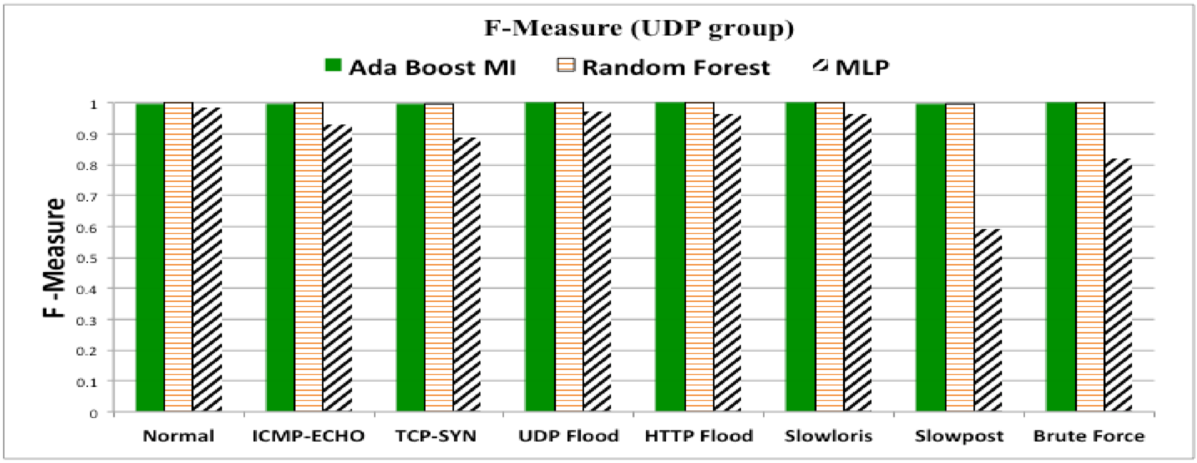
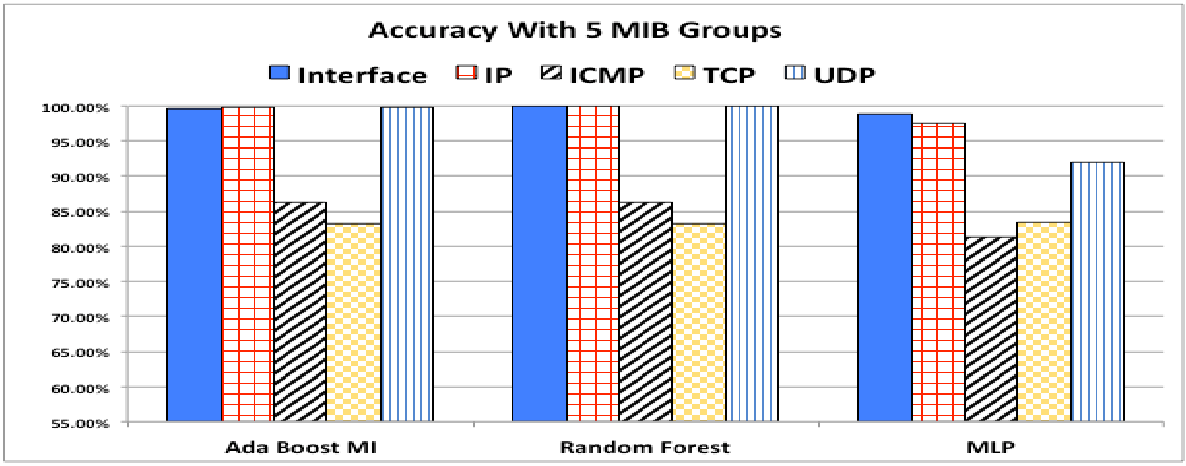


Figure 8 Accuracy rate for the three classifiers with the 5 MIB group datasets



6 Conclusions

In this paper, we have presented a methodology for network attack detection based on SNMP-MIB data by applying machine learning techniques. The purpose of our work was to prove the ability and effectiveness of SNMP-MIB data in network anomaly detection by demonstrating the detection of the largest possible number of the most common and modern attacks that can occur on different network layers (network layer, transport layer and application layer). Our methodology involved three classification algorithms, namely AdaboostM1, Random Forest and MLP classifiers. In our approach, we categorised the MIB variables in five MIB groups (Interface, IP, ICMP, TCP and UDP) where each group included a number of MIB variables that were affiliated to it. The classification algorithms were then applied to each MIB group separately, in order to show how each group is affected by attacks, and therefore to determine the most effective group(s) in anomaly detection. From the results of this approach, we found that the performance of each classifier is different over all the MIB groups, where the accuracy rate varied between high and low for the three classifiers. The Random Forest classifier achieved the highest accuracy rate with the IP group (100%) and with the Interface group (99.93%). From the results, we also found that among the five MIB groups the Interface and IP groups were the groups that were affected the most by all types of attack, while the ICMP, TCP and UDP groups were less affected. The overall conclusion is that using SNMP-MIB data with machine learning techniques is a very effective approach to network anomaly detection with a distinctive effect on network security.

References

Al-Kasassbeh, M. and Adda, M. (2009) 'Network fault detection with wiener filter-based agent', *Journal of Network and Computer Applications*, Vol. 32, No. 4, pp.824–833.

Al-Kasassbeh, M. (2011) 'Network intrusion detection with wiener filter-based agent', *World Applied Sciences Journal (WASJ)*, Vol. 13, No. 11, pp.2372–2384.

Al-Kasassbeh, M., Al-Naymat, G. and Al-Hawari, E. (2016) 'Towards generating realistic SNMP-MIB dataset for network anomaly detection', *International Journal of Computer Science and Information Security*, Vol. 14, No. 9, pp.1162–1185.

Al-Kasassbeh, M., Al-Naymat, G., Hassanat, A., Almseidin, M. (2016) 'Detecting distributed denial of service attacks using data mining techniques', *International Journal of Advanced Computer Science and Application*, Vol. 7, No. 1, pp.436–445.

Bao, C.M. (2009) 'Intrusion detection based on one-class svm and snmp mib data', *IEEE 5th International Conference on Information Assurance and Security*, Vol. 2, pp.346–349.

Cabrera, J.B., Lewis, L., Qin, X., Lee, W. and Mehra, R.K. (2002) 'Proactive intrusion detection and distributed denial of service attacks – a case study in security management', *Journal of Network and Systems Management*, Vol. 10, No. 2, pp.225–254.

Cerroni, W., Moro, G., Pirini, T. and Ramilli, M. (2013) 'Peer-to-peer data mining classifiers for decentralized detection of network attacks', *Proceedings of the 24th Australasian Database Conference*, Vol. 137, pp.101–107.

Cerroni, W., Moro, G., Pasolini, R. and Ramilli, M. (2015) 'Decentralized detection of network attacks through P2P data clustering of SNMP data', *Computers & Security*, Vol. 52, pp.1–16.

Gupta, J. and Singh, J. (2017) 'Detecting anomaly based network intrusion using feature extraction and classification techniques', *International Journal of Advanced Research in Computer Science*, Vol. 8, No. 5, pp.1453–1456

Hsiao, H.W., Lin, C.S. and Chang, S.Y. (2009) 'Constructing an ARP attack detection system with SNMP traffic data mining', *Proceedings of the 11th International Conference on Electronic Commerce (ICEC'09)*, ACM, New York, NY, USA, pp.341–345.

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P. and Witten, I.H. (2009) 'The WEKA data mining software: an update', *ACM SIGKDD Explorations Newsletter*, Vol. 11, No. 1, pp.10–18.

Kirnapure, W.K. and Patil, A.R.B. (2017) 'Classification, detection and prevention of network attacks using rule based approach', *International Research Journal of Engineering and Technology*, Vol. 4, No. 4, pp.1453–1459.

Li, J. and Manikopoulos, C. (2003) 'Early statistical anomaly intrusion detection of DOS attacks using MIB traffic parameters', *IEEE Systems, Man and Cybernetics Society Information Assurance Workshop*, pp.53–59.

Moore, D., Shannon, C., Brown, D.J., Voelker, G.M. and Savage, S. (2006) 'Inferring internet denial-of-service activity', *ACM Transactions on Computer Systems*, Vol. 24, No. 2, pp.115–139.

Muda, Z., Yassin, W., Sulaiman, M.N. and Udzir, N.I. (2016) 'K-means clustering and naive bayes classification for intrusion detection', *Journal of IT in Asia*, Vol. 4, No. 1, pp.13–25.

Nanda, N.B. and Parikh, A. (2017) 'Classification and technical analysis of network intrusion detection systems', *International Journal of Advanced Research in Computer Science*, Vol. 8, No. 5, pp.657–661.

Namvarasl, S. and Ahmadzadeh, M. (2014) 'A dynamic flooding attack detection system based on different classification techniques and using SNMP MIB data', *International Journal of Computer Networks and Communications Security*, Vol. 2, No. 9, pp.279–284.

Park, J.S. and Kim, M.S. (2008) 'Design and implementation of an SNMP-based traffic flooding attack detection system', in Ma, Y., Choi, D. and Ata, S. (Eds): *Challenges for Next Generation Network Operations and Service Management (APNOMS'08)*, Springer, Berlin, Heidelberg, Vol. 5297, pp.380–389.

Priya, P.M., Akilandeswari, V., Shalinie, S.M., Lavanya, V. and Priya, M.S. (2014) 'The protocol independent detection and classification (PIDC) system for DRDoS attack', *IEEE International Conference on Recent Trends in Information Technology (ICRTIT)*, pp.1–7.

Puttini, R., Hanashiro, M., Miziara, F., de Sousa, R., García-Villalba, L.J. and Barenco, C.J. (2006) 'On the anomaly intrusion-detection in mobile ad hoc network environments', in Cuenca, P. and Orozco-Barbosa, L. (Eds): *Personal Wireless Communications (PWC'06)*, Springer, Berlin, Heidelberg, Vol. 4217, pp.182–193.

Rahmani, C., Sharifi, M. and Tafazzoli, T. (2004) 'An experimental analysis of proactive detection of distributed denial of service attacks', *Proceedings of the IIT Kanpur Hacker's Workshop (IITKHACK04),* pp.37–44.

Thottan, M. and Ji, C. (2003) 'Anomaly detection in IP networks', *IEEE Transactions on Signal Processing*, Vol. 51, No. 8, pp.2191–2204.

Vyavhare, A., Bhosale, V., Sawant, M. and Girkar, F. (2012) 'Co-operative wireless wireless intrusion detection system using MIBs from SNMP, *International Journal of Network Security & Its Applications*, Vol. 4, No. 2, pp.147–154.

Wu, Q. and Shao, Z. (2005) 'Network anomaly detection using time series analysis', *Proceedings of the Joint International Conference on Autonomic and Autonomous Systems and International Conference on Networking and Services (ICAS-ICNS'05)*, IEEE Computer Society, Washington, DC, USA, pp. 42–51.

Yu, J., Kang, H., Park, D., Bang, H.C. and Kang, D.W. (2013) 'An in-depth analysis on traffic flooding attacks detection and system using data mining techniques', *Journal of Systems Architecture*, Vol. 59, No. 10, pp.1005–1012.

Yu, J., Lee, H., Kim, M.S. and Park, D. (2008) 'Traffic flooding attack detection with SNMP MIB using SVM', *Computer Communications*, Vol. 31, No. 17, pp.4212–4219.